

The Catalan Numbers
Algebra Seminar — 02/12/2026
Ahmer Nadeem Khan

Summary

- **Context:** A probability question from a finance interview led me to the Catalan numbers; I first encountered them while reading Stanley's *Enumerative Combinatorics*, and later June Huh recommended Stanley's book devoted specifically to the Catalan numbers.
- **Main statement:** The n th Catalan number is

$$C_n = \frac{1}{n+1} \binom{2n}{n} \quad (n \geq 0).$$

and it counts many families of combinatorial objects, for example:

1. Polygon triangulations
 2. Dyck paths
 3. Binary trees
 4. Plane trees
 5. Ballot sequences
 6. Parenthesizations
- **Key idea:** Fundamental recurrence; various bijections.

$$C_0 = 1, \quad C_{n+1} = \sum_{k=0}^n C_k C_{n-k} \quad (n \geq 0).$$

- **Takeaway** The Catalan numbers are ubiquitous in combinatorics, and count various useful and interesting objects.

Roadmap

1. "Green Book" probability problem
2. Richard P. Stanley's book *The Catalan Numbers* (recommended by June Huh)
3. Definitions: formal power series; polygon triangulations
4. Proof of Catalan numbers via generating functions
5. Elegant proof via reflection principle
6. Main bijections
7. Other Catalan-type objects; generalizations
8. Importance of certain objects

“Green Book” probability problem

Theater tickets (probability / ballot interpretation)

Problem. A theater sells tickets for \$5, but starts with no change. There are $2n$ customers in a line; exactly n customers have a \$5 bill and n customers have a \$10 bill, and the order of the line is uniformly random among all such sequences. What is the probability that the theater can sell tickets to the customers in order (with no rearranging)?

- Problem in a Quantitative Finance interview book.
- The problem describes well known sequences called ballot sequences with a rich history (one of the first objects to be studied).
- Explain why they are called ballot sequences (candidates A and B with n votes each such that A is never trailing B).
- Bijection with Dyck paths (proof via the reflection principle).
- Generalizations to Fuss–Catalan numbers.

$$\#A \geq k \cdot \#B.$$

$(kn+2)$ -gon int
 $(k+2)$ -gons

Richard P. Stanley’s book *The Catalan Numbers* (recommended by June Huh)

- A monograph on the Catalan numbers.
- Contains most of the results discussed.
- Contains a history of the Catalan numbers.
- Also contains 214 different combinatorial interpretations from a list evolving over 40 years.
- Contains many problems as well.

Definitions: formal power series; polygon triangulations

$\mathbb{C}[[x]]$ (formal power series)

$$\mathbb{C}[[x]] := \left\{ \sum_{n \geq 0} a_n x^n : a_n \in \mathbb{C} \right\}.$$

Generating functions

Multiplication (Cauchy product):

$$\left(\sum_{n \geq 0} a_n x^n \right) \left(\sum_{n \geq 0} b_n x^n \right) = \sum_{n \geq 0} \left(\sum_{k=0}^n a_k b_{n-k} \right) x^n.$$

Chain of ideals (powers of x):

$$(x) \supset (x^2) \supset (x^3) \supset \dots \quad \text{where} \quad (x^m) = \{x^m g(x) : g(x) \in \mathbb{C}[[x]]\}.$$

Why PID (one sentence): every nonzero ideal I contains an element of minimal x -order k , and then $x^k \in I$, so $I = (x^k)$.

Hence DVR: $\mathbb{C}[[x]]$ is local with unique maximal ideal (x) , and every $0 \neq f \in \mathbb{C}[[x]]$ factors uniquely as

$$f = x^{v(f)} u, \quad u \in \mathbb{C}[[x]]^\times, \quad v(f) = \min\{n : a_n \neq 0\}.$$

x -adic topology: For $F(x) = \sum_{n \geq 0} a_n x^n$, define

$$\deg(F) = \min\{n : a_n \neq 0\}, \quad \deg(0) = \infty.$$

A sequence $F_i(x)$ converges to $F(x)$ iff

$$\deg(F - F_i) \rightarrow \infty.$$

metrizable / complete

Same coset mod (x^N) / same ideal condition:

$$f \equiv g \pmod{x^N} \iff f - g \in (x^N) \iff f \text{ and } g \text{ agree in coefficients of } x^0, \dots, x^{N-1}.$$

- Can develop formal calculus.
- Define sum, product, compositions, differentiation, and convergence (complete).
- Geometric series:

$$\left(\sum_{n \geq 0} \alpha^n x^n \right) (1 - \alpha x) = \sum_{n \geq 0} c_n x^n,$$

$$c_n = \begin{cases} 1, & n = 0, \\ \alpha^n - \alpha(\alpha^{n-1}) = 0, & n \geq 1. \end{cases}$$

- Fibonacci sequence:

$$\sum_{n \geq 0} f(n) \frac{x^n}{n!} = \exp\left(x + \frac{x^2}{2}\right)$$

differentiating

$$\sum_{n \geq 1} f(n) \frac{x^{n-1}}{(n-1)!} = (1+x)e^{x+x^2/2} = (1+x) \sum_{n \geq 0} f(n) \frac{x^n}{n!}.$$

Equating coefficients of $x^n/n!$ yields

$$f(n+1) = f(n) + nf(n-1), \quad n \geq 1.$$

$e^x e^{-x} = 1$

$$\sum_{n \geq 0} \frac{x^n}{n!} \sum_{n \geq 0} \frac{x^n}{n!} (-1)^n = 1$$

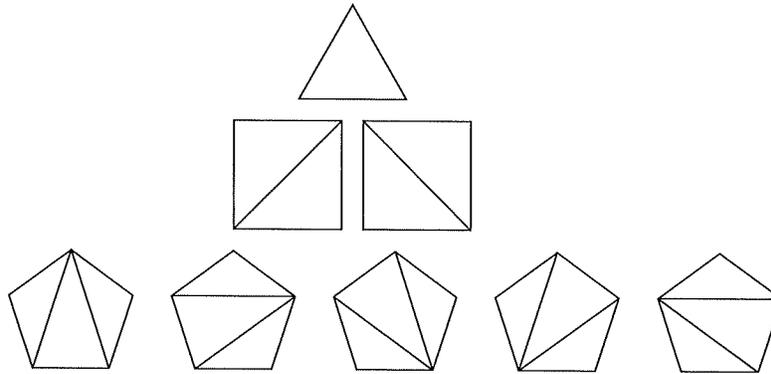
you can go the other way.

Polygon triangulations

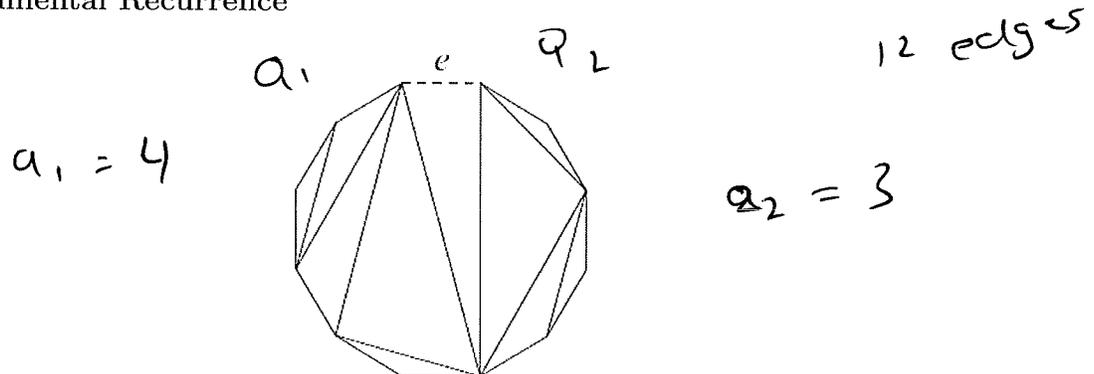
Let P_{n+2} be a convex polygon with $n + 2$ vertices. A triangulation is a collection of non-crossing diagonals that partition the polygon into n triangles using $n - 1$ diagonals.

Define the n th Catalan number C_n to be the number of triangulations of P_{n+2} .

$$C_0 = 1, \quad C_1 = 1, \quad C_2 = 2, \quad C_3 = 5, \quad C_4 = 14, \dots C_{10} = 16796.$$



The Fundamental Recurrence



Let P_{n+3} be a convex $(n + 3)$ -gon and fix an edge e . For a triangulation T , removing the triangle containing e produces two triangulated polygons Q_1, Q_2 with one common vertex.

If Q_i has $a_i + 2$ vertices, then

$$a_1 + a_2 = n.$$

Edge case: if the triangle containing e uses a boundary edge adjacent to e , then one Q_i is a single edge; treat it as a 2-gon with

$$C_0 = 1.$$

Independent triangulations give $C_{a_1}C_{a_2}$, and conversely, any triangulations of Q_1, Q_2 glue uniquely along e to form a triangulation of P_{n+3} .

$$C_{n+1} = \sum_{k=0}^n C_k C_{n-k}, \quad C_0 = 1. \quad (\text{almost impossible sometimes})$$

Proof of Catalan numbers via generating functions

Theorem 1 (Generalized Binomial Theorem, due to Isaac Newton). For any $a \in \mathbb{C}$,

$$(1+x)^a = \sum_{n \geq 0} \binom{a}{n} x^n,$$

where

$$\binom{a}{n} = \frac{a(a-1)\cdots(a-n+1)}{n!}, \quad \binom{a}{0} = 1.$$

- Derived from Taylor series of $(1+x)^a$ at $x=0$. Justified through the calculus we mentioned before, and other ways to prove it also exist.

Proposition 2. Let

$$\begin{aligned} C(x) &= \sum_{n \geq 0} C_n x^n \\ &= 1 + x + 2x^2 + 5x^3 + 14x^4 + 42x^5 + 132x^6 + 429x^7 + 1430x^8 + \dots \end{aligned}$$

Then

$$C(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

Proof. Multiply $C_{n+1} = \sum_{k=0}^n C_k C_{n-k}$ by x^n and sum over $n \geq 0$:

$$\sum_{n \geq 0} C_{n+1} x^n = \frac{C(x)-1}{x}, \quad \sum_{n \geq 0} \left(\sum_{k=0}^n C_k C_{n-k} \right) x^n = C(x)^2.$$

Hence

$$\frac{C(x)-1}{x} = C(x)^2 \Rightarrow xC(x)^2 - C(x) + 1 = 0.$$

Solving,

$$C(x) = \frac{1 \pm \sqrt{1-4x}}{2x}.$$

Since $\sqrt{1-4x} = 1 - 2x + \dots$, the $+$ branch gives $\frac{1+(1-2x+\dots)}{2x} = x^{-1} - 1 + \dots$ (not a power series). Therefore

$$C(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

□

Theorem 3. For $n \geq 0$,

$$C_n = \frac{1}{n+1} \binom{2n}{n} = \frac{(2n)!}{n!(n+1)!}.$$

Proof. From

$$C(x) = \frac{1 - \sqrt{1-4x}}{2x},$$

$-1 \rightarrow \sum_{n \geq 0} C_n x^n$
 $1/x \rightarrow \sum_{n \geq 1} C_n x^{n-1}$
 \otimes
 $k = n-1$
 $\sum_{n \geq 0} C_{n+1} x^n$

← Cauchy product

use the generalized binomial theorem:

$$\sqrt{1-4x} = (1-4x)^{1/2} = \sum_{n \geq 0} \binom{1/2}{n} (-4x)^n.$$

Hence

$$C(x) = \frac{1}{2x} \left(1 - \sum_{n \geq 0} \binom{1/2}{n} (-4)^n x^n \right) = -\frac{1}{2} \sum_{n \geq 0} \binom{1/2}{n+1} (-4)^{n+1} x^n.$$

Equating coefficients of x^n gives

$$C_n = -\frac{1}{2} \binom{1/2}{n+1} (-4)^{n+1}.$$

A routine simplification of the RHS yields $C_n = \frac{1}{n+1} \binom{2n}{n}$.

$$= \frac{2n!}{n!n!}$$

□

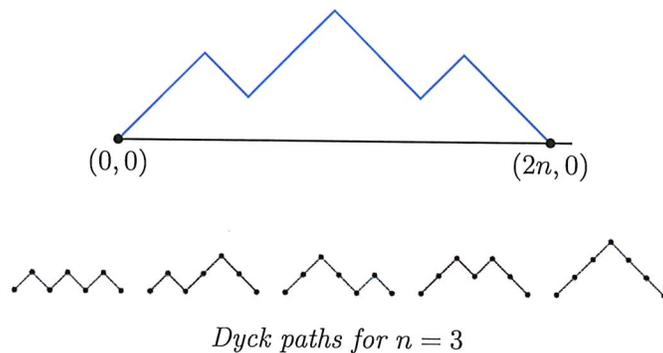
$$\begin{aligned} & \frac{1}{2x} \sum_{n \geq 0} \binom{1/2}{n} (-4)^n x^n \\ &= -\frac{1}{2x} \left(\sum_{n \geq 1} \binom{1/2}{n} (-4)^n x^n \right) \\ &= -\frac{1}{2x} \left(\sum_{n \geq 0} \binom{1/2}{n+1} (-4)^{n+1} x^{n+1} \right) \\ &= -\frac{1}{2} \binom{1/2}{n+1} (-4)^{n+1} x^{\frac{1}{2}-2n} \\ &= -\frac{1}{2} \left(\frac{\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2} (-1)^{\frac{2n-1}{2}} (\frac{1}{2} - (n+1) + 1)!}{(n+1)!} \right) (-2)^{2n+1} \\ &= \frac{1}{n!(n+1)} \frac{(-1/2)(-3/2)(-5/2) \cdots (-1 \frac{(2n-1)}{2})}{(n+1)!} (-2)^{2n+1} \\ &= \frac{1}{n!(n+1)} \end{aligned}$$

Elegant proof via reflection principle

- The original problem is a problem of ballot sequences.
- There is an obvious bijection between Dyck paths.

Let $S \subseteq \mathbb{Z}^d$. A **lattice path** in \mathbb{Z}^d of length k with steps in S is a sequence $v_0, v_1, \dots, v_k \in \mathbb{Z}^d$ such that $v_i - v_{i-1} \in S$ for each i . We say the path goes from v_0 to v_k .

A **Dyck path** of length $2n$ (or semi-length n) is a lattice path in \mathbb{Z}^2 from $(0,0)$ to $(2n,0)$ with steps $(1,1)$ and $(1,-1)$ that never passes below the x -axis.



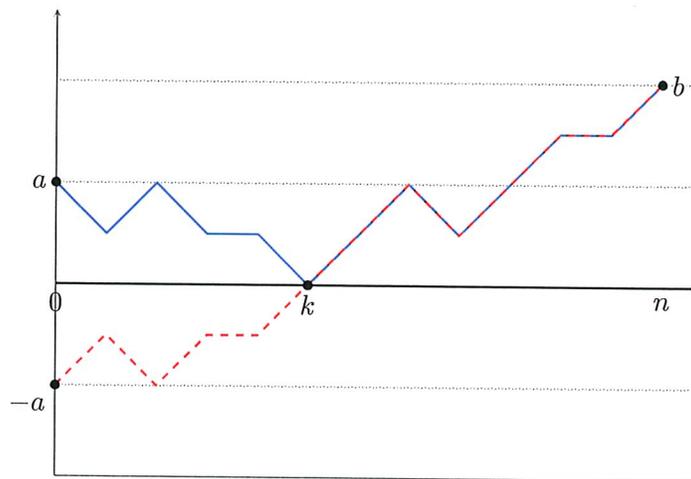
The reflection principle (random walks)

Theorem 4 (Reflection principle). Consider lattice paths in \mathbb{Z}^2 consisting of steps $(1,+1)$ and $(1,-1)$.

Fix $n \geq 1$ and $a \geq 1$. Let $N_n(a,b)$ denote the number of paths of length n starting at height a and ending at height b . Let $N_n^0(a,b)$ denote the number of such paths that hit height 0 at least once i.e., there exists $0 < k < n$ such that $v_k = (k,0)$.

Then

$$N_n^0(a,b) = N_n(-a,b).$$



Ballot problem via reflection.

Encode the line as a lattice path with steps $+1$ (a \$5 bill) and -1 (a \$10 bill). Paths go from $(0, 0)$ to $(2n, 0)$, so there are

$$\binom{2n}{n}$$

total paths.

We count paths that never go below the x -axis. By the reflection principle, paths that hit -1 correspond bijectively to paths ending at $(2n, -2)$, which must have $(n-1)$ up-steps and $(n+1)$ down-steps:

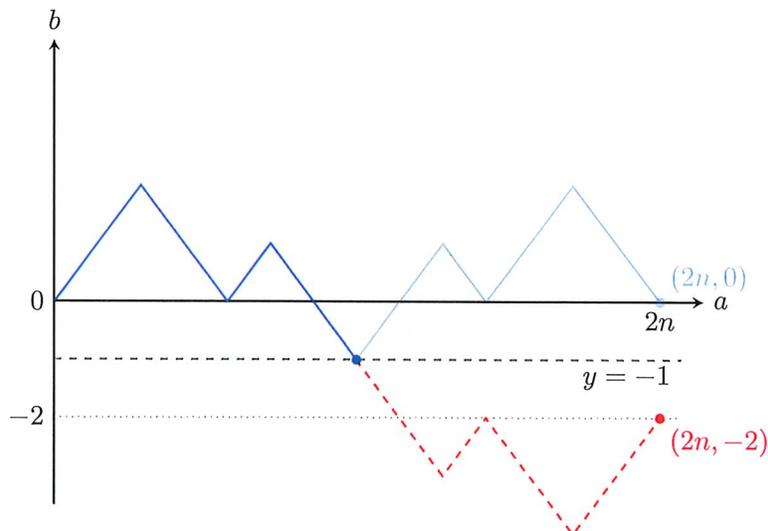
$$\binom{2n}{n-1}.$$

Hence the number of Dyck paths is

$$\binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{n+1} \binom{2n}{n}.$$

Therefore the probability that no one must change positions is

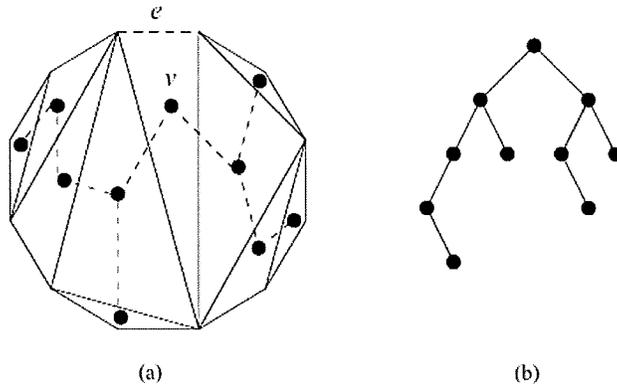
$$\frac{1}{n+1}.$$



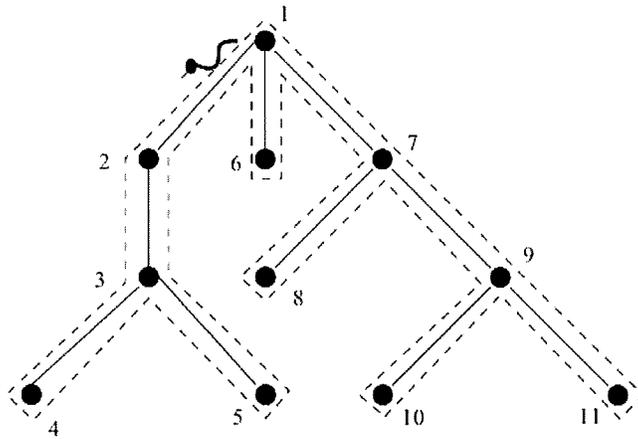
Main bijections

Theorem 5 (Catalan interpretations). *The Catalan number C_n counts the following:*

1. *Triangulations T of a convex polygon with $n + 2$ vertices.*
2. *Binary trees B with n vertices.*
3. *Plane trees P with $n + 1$ vertices.*
4. *Ballot sequences of length $2n$.*
5. *Parenthesizations (or bracketings) of a string of $n + 1$ x 's subject to a nonassociative binary operation.*
6. *Dyck paths of length $2n$.*



A binary tree associated with a triangulated polygon

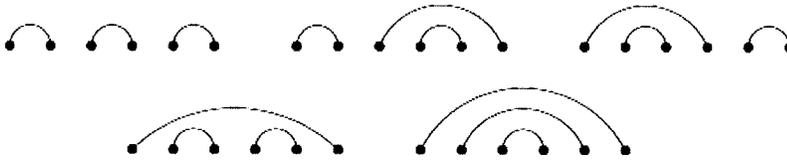


Preorder on a plane tree

Other Catalan-type objects; generalizations

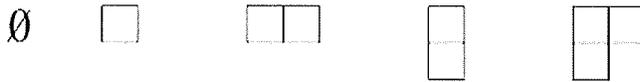
- *The Catalan Numbers* contains a list of 214 different combinatorial interpretations.
- Complete binary trees (that is, binary trees such that every vertex has zero or two children) with $2n + 1$ vertices (or $n + 1$ endpoints).
- Plane trees with $n + 2$ vertices such that the rightmost path of each subtree of the root has even length.

61. *Noncrossing (complete) matchings* on $2n$ vertices, i.e., ways of connecting $2n$ points in the plane lying on a horizontal line by n nonintersecting arcs, each arc connecting two of the points and lying above the points.



- Dyck paths of length $2n + 2$ whose first down step is followed by another down step.
- Peaks of height one (or hills) in all Dyck paths from $(0, 0)$ to $(2n, 0)$.

167. Young diagrams that fit in the shape $(n - 1, n - 2, \dots, 1)$.



168. Standard Young tableaux of shape (n, n) (or equivalently, of shape $(n, n - 1)$).

123 124 125 134 135
456 356 346 256 246

or

123 124 125 134 135
45 35 34 25 24

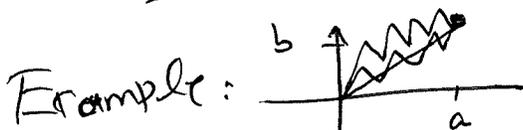
fill 1 through n , increasing strictly in row and column.

• Fuss-Catalan

convex $(k+2)$ -gon into

$(k+2)$ -gons $C_{n,k}$

• (a, b) -Catalan



11
$$\text{Cat}(a, b) = \frac{1}{a+b} \binom{a+b}{a}$$

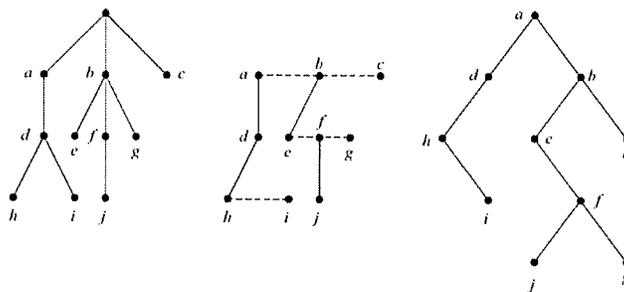
• Motzkin Numbers : $(0, 0) \rightarrow (n, 0)$

$$S = (1, 0), (1, 1), (1, -1)$$

• Super-Catalan Numbers : $S(m, n) = \frac{(2m)!(2n)!}{m!n!(m+n)!}$

↳ ^{stated} ~~proved~~ by Catalan
to be integers

↳ Open problem [Combinatorial Interpretation]



Binary to plane tree (due to Bruijn and Morselt)

- Knuth calls this the "natural correspondence".

Recurrence for parenthesizing.

Fix a full bracketing of $x_0x_1 \cdots x_n$ and look at the *last* operation: it has the form

$$(\text{bracketing of } x_0 \cdots x_k) \star (\text{bracketing of } x_{k+1} \cdots x_n)$$

for a unique $k \in \{0, \dots, n-1\}$. Thus for fixed k there are $C_k \cdot C_{n-1-k}$ choices, so

$$C_n = \sum_{k=0}^{n-1} C_k C_{n-1-k} \quad (C_0 = 1).$$

Recurrence via ballot sequences.

Let $\beta = a_1 \cdots a_{2n+2}$ be a ballot sequence and let

$$k = \min\{j \geq 0 : a_1 + \cdots + a_{2j+2} = 0\}.$$

Then β decomposes uniquely into ballot sequences: $a_2 \cdots a_{2k+1}$ (of length $2k$) and $a_{2k+3} \cdots a_{2n+2}$ (of length $2(n-k)$), giving $C_k C_{n-k}$ possibilities.

Importance of certain objects

Coxeter Group $\langle r_1, r_2, \dots, r_n \mid (r_i r_j)^{m_{ij}} = 1 \rangle$
 i.e. $r_i^2 = 1 \ \forall i$ (involutions) $[m_{ii} = 1]$
 and $m_{ij} = m_{ji} \geq 2 \ i \neq j$

Coxeter length: $l(v) =$ ^{min.} length of a reduced word of v

Bruhat order: $u \leq w$ iff a reduced word for u appears as a subword of some reduced word for w .

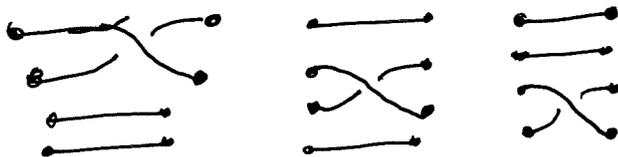
$\rightarrow l(w)$ in S_n
 $= \#$ of inversions

$s_i = (i \ i+1)$ generate S_n (type A_{n-1})
 $S_n = \langle s_1, \dots, s_{n-1} \mid s_i^2 = 1, s_i s_j = s_j s_i \ | \ |i-j| > 1$
 $s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1} \rangle$

Matsumoto's Theorem

If 2 reduced words represent the same element, then the 1st \leftrightarrow 2nd by $xyxy \dots$ to $yxyx \dots$

Braid Groups:



Generators g_1, g_2, g_3 of B_4

$$B_n = \langle g_1, \dots, g_{n-1} \mid g_i g_{i+1} g_i = g_{i+1} g_i g_{i+1}, g_i g_j = g_j g_i \ | \ |i-j| \geq 2 \rangle$$

Artin Groups $stst \dots = tst \dots$
 (Artin relations)₁₂

Hecke Algebra / Kazhdan-Lusztig Polynomials / (Deligne)

Temperley-Lieb Algebra

Intersection cohomology groups of

Jones relations!

$$E_i^2 = E_i, \beta E_i E_j E_i = E_i |i-j| \neq 0$$

$$E_i E_j = E_j E_i \quad |i-j| \geq 2$$

$$\beta \in K \quad \dim(TL_n(\beta)) = C_n$$

Schubert varieties

$$\left[\begin{array}{l} \text{Proj. deg} \\ = \# \text{ Gr}(k, n) \end{array} \right]$$

→ degree of ~~the~~ $Gr(2n+2) = C_n$

∩ { general linear subspace of complementary dimension }

↳ # SYT of $k \times (n-k)$ shape

$$Gr(k, n)$$

$$\mapsto P(A^k \mathbb{C}^n)$$

↳ Hook-length formula!

$$Gr(k, n) \subset \mathbb{P}^N \quad N = \binom{n}{k} - 1$$

Plücker embedding

$$\dim(Gr(k, n)) = k(n-k)$$

Schubert Calculus (Pattern Avoidance)

$$[X_u] \cup [X_v] = \sum C_{uv} [X_w]$$

$$F_\bullet = (F_1, \dots, F_n) \in \mathbb{C}^n \quad \dim(F_i) = i \quad (f_1, \dots, f_n)$$

$$E_\bullet = (E_1, \dots, E_n) \quad Fl_n(\mathbb{C}) \cong S_n(\mathbb{C}) / B$$

position $(E_\bullet, F_\bullet) =$ permutation matrix in canonical form

$$F_\bullet = \langle 2e_1, e_2, 2e_1 + e_3, 2e_1 + e_4, e_1 \rangle \approx \begin{pmatrix} 2 & 2 & 7 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

For $w \in S_n$

Schubert cell $C_w(E_\bullet) \subset Fl_n(\mathbb{C}) =$ all flags F_\bullet

$$\dim(C_w(\mathbb{C})) = l(w) \quad \text{s.t. } (E_\bullet, F_\bullet) = w$$

Schubert variety

$$C_{2311} = \left\{ \begin{pmatrix} x & y & z & 1 \\ 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix} : x, y, z \in \mathbb{C} \right\}$$

$$X_w(E_\bullet) = \overline{C_w(E_\bullet)} \text{ in Zariski Topology}$$

$$X_w = \bigcup_{v \in w} C_v \quad \text{i.e.} \quad X_v \subset X_w$$

Poincaré Polynomials
for $H^*(X_w)$

$$= P_w(t) = \sum_{v \in w} t^{2\ell(v)}$$

C_n = multiplicity of
 X_{w_0} in Ω_w

iff $v \in w$ (Bruhat
order)

$$w_0 = n \ n-1 \ \dots \ 1$$

$$X_{w_0}$$

$$w = n \ 2, 3 \ \dots, n-2, n-1, n$$

⑧ X_w is singular iff w is 3412 and 4231
smooth avoiding.

→ many other Pattern Avoidance Properties

⑧ - Lakshmibai - Sandhya Theorem!

□

References (quick pointers)

1. Richard P. Stanley, *The Catalan Numbers*.
2. Richard P. Stanley, *Enumerative Combinatorics, Volume I*.
3. Xinfeng Zhou, *A Practical Guide to Quantitative Finance Interviews*.
4. Haraku Abe, Sarah Billey, *Consequences of the Lakshmibai-Sandhya Theorem: the ubiquity of permutation patterns in Schubert calculus and related geometry*. 2010 Mathematics Subject Classification: 14M15, 14B05, 20F55.
5. Various Wikipedia entries e.g. Bruhat Order, Braid Group, Coxeter Group, Artin–Tits group, Matsumoto’s theorem (group theory), Iwahori–Hecke algebra, Temperley–Lieb algebra, Kazhdan–Lusztig polynomial, Cartan–Dieudonné theorem.